































































































































































- BLOOM, H. S. (2005): *Randomizing groups to evaluate place-based programs* NY: Russell Sage Foundation, chap. Learning more from social experiments, pp. 115–172.
- BOBONIS, G. J., E. MIGUEL, AND C. P. SHARMA (2004): “Iron Deficiency Anemia and School Participation,” Poverty Action Lab Paper No. 7.
- BUDDLEMEYER, H., AND E. SKOFIAS (2003): “An Evaluation on the Performance of Regression Discontinuity Design on PROGRESA,” Institute for Study of Labor, Discussion Paper No. 827.
- CAMPBELL, D. T. (1969): “Reforms as Experiments,” *American Psychologist*, 24, 407–429.
- CARD, D. (1999): “The Causal Effect of Education on Earnings,” in *Handbook of Labor Economics*, Vol. 3A, ed. by O. Ashenfelter, and D. Card. North Holland, pp. 1801–63.
- CARD, D., AND A. B. KRUEGER (1995): *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton, NJ: Princeton University Press.
- CHATTOPADHYAY, R., AND E. DUFLO (2004): “Women as Policy Makers: Evidence from a Randomized Policy Experiment in India,” *Econometrica*, 72(5), 1409–1443.
- COHEN, J. (1988): *Statistical Power Analysis for the Behavioral Science*. Hillsdale, NJ: Lawrence Erlbaum, 2nd edition edn.
- COOK, T. D., W. R. SHADISH, AND V. C. WONG (2006): “Within Study Comparisons of Experiments and Non-Experiments: Can they help decide on Evaluation Policy,” mimeo, Northwestern University.
- COX, D., AND N. REID (2000): *Theory of the Design of Experiments*. London: Chapman and Hall.
- DAS, J., P. KRISHNAN, J. HABYARIMANA, AND S. DERCON (2004): “When can school inputs improve test scores?,” World Bank Policy Research working paper; no. WPS 3217.
- DEATON, A. (1997): *The Analysis of Household Surveys*. World Bank, International Bank for Reconstruction and Development.

- DEHEJIA, R. H., AND S. WAHBA (1999): “Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs,” *Journal of the American Statistical Association*, 94(448), 1053–62.
- DELONG, J. B., AND K. LANG (1992): “Are All Economic Hypotheses False?,” *Journal of Political Economy*, 100:6, 1257–72.
- DIAZ, J. J., AND S. HANDA (2006): “An Assessment of Propensity Score Matching as a Non Experimental Impact Estimator: Evidence from Mexico’s Progresa Program,” forthcoming, *Journal of Human Resources*.
- DICKSON, RUMONA, S. A. P. W. C. D., AND P. GARNER (2000): “Effect of treatment for intestinal helminth infection on growth and cognitive performance in children: systematic review of randomized trials,” *British Medical Journal*, 320, 1697–1701.
- DONALD, S., AND K. LANG (2001): “Inference with Differences-in-Differences and Other Panel Data,” Discussion paper, Boston University Department of Economics.
- DUFLO, E. (2004): *Scaling Up and Evaluation* Oxford University Press and World Bank, chap. Accelerating Development.
- (2006): “Field Experiments in Development Economics,” Discussion paper.
- DUFLO, E., P. DUPAS, M. KREMER, AND S. SINEI (2006): “Education and HIV/AIDS Prevention: Evidence from a randomized evaluation in Western Kenya,” mimeo, MIT.
- DUFLO, E., AND R. HANNA (2006): “Monitoring Works: Getting Teachers to Come to School,” NBER Working Paper No. 11880.
- DUFLO, E., AND M. KREMER (2005): “Use of Randomization in the Evaluation of Development Effectiveness,” in *Evaluating Development Effectiveness*, ed. by O. Feinstein, G. K. Ingram, and G. K. Pitman. New Brunswick, New Jersey and London, U.K.: Transaction Publishers, vol. 7, pp. 205–232.
- DUFLO, E., M. KREMER, AND J. ROBINSON (2006): “Understanding Technology Adoption: Fertilizer in Western Kenya, Preliminary Results from Field Experiments,” mimeo.

- DUFLO, E., S. MULLAINATHAN, AND M. BERTRAND (2004): “How Much Should We Trust Difference in Differences Estimates?,” *Quarterly Journal of Economics*, 119(1), 249–275.
- DUFLO, E., AND E. SAEZ (2003): “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment,” *Quarterly Journal of Economics* 118(3), 2003, pp. 815–842, 118(3), 815–842.
- DUPAS, P. (2006): “Relative Risks and the Market for Sex: Teenagers, Sugar Daddies, and HIV in Kenya,” mimeo, Dartmouth College.
- FISHER, R. A. (1926): “The Arrangement of Field Experiments,” *Journal of the Ministry of Agriculture*, 33, 503–513.
- GERTLER, P. J., AND S. BOYCE (2001): “An Experiment in Incentive-Based Welfare: The Impact of PROGRESA on Health in Mexico,” Mimeo, UC-Berkeley.
- GLAZERMAN, S., D. LEVY, AND D. MYERS (2003): *Nonexperimental Replications of Social Experiments: A Systematic Review*. Princeton, NJ: Mathematica Policy Research, Inc.
- GLEWWE, PAUL, N. I., AND M. KREMER (2003): “Teacher Incentives,” Working Paper 9671, National Bureau of Economic Research.
- GLEWWE, P., KREMER, AND S. MOULIN (2004): “Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya,” Mimeo, Harvard University.
- GLEWWE, P., AND M. KREMER (2005): “Schools, Teachers, and Education Outcomes in Developing Countries,” *Handbook on the Economics of Education* (forthcoming).
- GLEWWE, P., M. KREMER, S. MOULIN, AND E. ZITZEWITZ (2004): “Retrospective Vs. Prospective Analyses of School Inputs: The Case of Flip Charts in Kenya,” *Journal of Development Economics*, 74, 251–268.
- GRASDAL, A. (2001): “The performance of sample selection estimators to control for attrition bias,” *Health Economics*, 10, 385–398.
- HARRISON, G., AND J. A. LIST (2004): “Field Experiments,” *Journal of Economic Literature*, XLII, 1013–1059.

- HAUSMAN, J. A., AND D. A. WISE (1979): "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment," *Econometrica*, 47(2), 455–73.
- HECKMAN, J., H. ICHIMURA, J. SMITH, AND P. TODD (1998): "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 66(5), 1017–1098.
- HECKMAN, J., H. ICHIMURA, AND P. TODD (1997): "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, 64(4), 605–54.
- (1998): "Matching as an Econometric Evaluation Estimator," *Review of Economic Studies*, 65(2), 261–94.
- HECKMAN, J., AND VYTLACIL (2005): "Structural Equations, Treatment Effects and Econometric Policy Evaluation," *Econometrica*, 73(3), 669–738.
- HECKMAN, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–161.
- HECKMAN, J. J., L. LOCHNER, AND C. TABER (1998): "General Equilibrium Treatment Effects: A Study of Tuition Policy," Working Paper 6426, National Bureau of Economic Research.
- HEDGES, L. (1992): "Modeling publication selection effects in meta-analysis," *Statistical Science*, 7, 227–236.
- HEDGES, L. V., AND I. OLKIN (1985): *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.
- HOLLAND, P. W. (1988): "Causal Inference, Path Analysis, and Recursive Structural Equations Models," *Sociological Methodology*, 18, 449–484.
- HOLM (1979): "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, 6, 65–70.
- HOXBY, C. M. (2003): "School Choice and School Productivity (Or, Could School Choice be a Rising Tide that Lifts All Boats)," Chicago: University of Chicago Press, chap. The Economics of School Choice.

- HSIEH, C.-T., AND M. S. URQUIOLA (2003): “When Schools Compete, How Do They Compete? An Assessment of Chile’s Nationwide School Voucher Program,” NBER Working Paper No. W10008.
- IMBENS, G., G. KING, AND G. RIDDER (2006): “On the Benefits of Stratification in Randomized Experiments,” Mimeo, Harvard.
- IMBENS, G. W. (2004): “Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review,” *The Review of Economics and Statistics*, 86(1), 4–29.
- INTERNATIONAL FOOD POLICY RESEARCH (2000): “Monitoring and Evaluation System,” Discussion paper, International Food Policy Research (IFPRI).
- KARLAN, AND J. ZINMAN (2005a): “Elasticities of Demand for Consumer Credit,” mimeo, Yale University.
- (2006a): “Expanding Credit Access: Using Randomized Supply Decisions To Estimate the Impacts,” mimeo, Yale University.
- KARLAN, D., AND J. ZINMAN (2006b): “Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment,” mimeo.
- KARLAN, D. S., AND J. ZINMAN (2005b): “Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment,” Working Papers 911, Economic Growth Center, Yale University, Available at <http://ideas.repec.org/p/egc/wpaper/911.html>.
- (2005c): “Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment,” Discussion paper, type: Working Papers; note: Available at <http://ideas.repec.org/p/egc/wpaper/911.html>.
- KELLER, S., AND B. SCHWARTZ (2001): “Final Evaluation Report: Contracting for Health Services Pilot Project,” Unpublished Asian Development Bank Report on Loan No. 1447-CAM.
- KLING, AND LIEBMAN (2004): “Experimental Analysis of Neighborhood Effects on Youth,” KSG Working Paper No. RWP04-034.

- KLING, J. R., J. B. LIEBMAN, L. F. KATZ, AND L. SANBONMATSU (2004): “Moving To Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-sufficiency and Health from a Randomized Housing Voucher Experiment,” .
- KREMER, M. (2003): “Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons,” *American Economic Review*, 93(2), 102–106.
- KREMER, M., E. MIGUEL, AND R. THORNTON (2004): “Incentives to Learn,” mimeo, Harvard University.
- KRUEGER, A., AND D. WHITMORE (2002): “Would smaller classes help close the blackwhite achievement gap?,” in *Bridging the achievement gap*, ed. by J. Chubb, and T. Loveless. Washington, DC: Brookings Institute Press, chap. Bridging the achievement gap.
- LALONDE, R. J. (1986): “Evaluating the Econometric Evaluations of Training Programs Using Experimental Data,” *American Economic Review*, 76(4), 602–620.
- LEE, D. S. (2002): “Trimming for Bounds on Treatment Effects with Missing Outcomes,” Working Paper 51.
- LIEBMAN, J. B., L. F. KATZ, AND J. KLING (2004): “Beyond Treatment Effects: Estimating the Relationship between Neighborhood Poverty and Individual Outcomes in the MTO Experiment,” Princeton IRS Working Paper 493.
- LOGAN, AND TAMHANE (2003): “Accurate critical constants for the one-sided approximate likelihood ratio test of a normal mean vector when the covariance matrix is estimated,” *Biometrics*, 58, 650–656.
- MALUCCIO, J. A., AND R. FLORES (2005): “Impact Evaluation of a Conditional Cash Transfer Program,” Discussion paper, International Food Policy Research Institute, Research Report No. 141.
- MANSKI, C. (1993): “Identification of Exogenous Social Effects: The Reflection Problem,” *Review of Economic Studies*, 60, 531–542.
- MANSKI, C. F. (1989): “Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon,” *Economics of Education Review*, 8(4), 305–312.

- MIGUEL, E., AND M. KREMER (2003): “Networks, Social Learning, and Technology Adoption: The Case of Deworming Drugs in Kenya,” Working Paper 61.
- (2004): “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities,” *Econometrica*, 72(1), 159–218.
- MOOKHERJEE, D. (2005): “Is There Too Little Theory in Development Economics?,” *Economic and Political Weekly*.
- MORDUCH, J. (1998): “Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh,” Mimeo, Princeton University.
- MOULTON, B. R. (1990): “An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units,” *Review of Economics and Statistics*, 72(2), 334–338.
- O’BRIEN (1984): “Procedures for Comparing Samples with Multiple Endpoints,” *Biometrics*, 40, 1079–1087.
- OLKEN, B. (2005): “Monitoring Corruption: Evidence from a Field Experiment in Indonesia,” mimeo, Harvard University.
- PARKER, S., L. RUBALCAVA, AND G. TERUEL (2005): “Evaluating Conditional Schooling-Health Transfer Programs (PROGRESA Program),” Forthcoming in Handbook of Development Economics, Volume 4.
- PITT, M., AND S. KHANDKER (1998): “The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter,” *Journal of Political Economy*.
- RAUDENBUSH, S. W., J. SPYBROOK, X. LIU, AND R. CONGDON (2005): “Optimal Design for Longitudinal and Multilevel Research: Documentation for the Optimal Design Software,” Retrieved April 15, 2005 from [http://sitemaker.umich.edu/group-based/optimal\\_design\\_software](http://sitemaker.umich.edu/group-based/optimal_design_software).
- RAVALLION, M. (2006): “Evaluating Anti-Poverty Programs,” forthcoming in Handbook of Development Economics Volume 4, edited by Robert E. Evenson and T. Paul Schultz, Amsterdam: North-Holland.

- ROSENBAUM, P., AND D. B. RUBIN (1983): “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70, 41–55.
- ROSENBAUM, P. R. (2002): “Covariance adjustment in randomized experiments and observational studies (with discussion),” *Statistical Science*, 17, 286327.
- RUBIN (1974): “Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies,” *Journal of Educational Psychology*, 66, 688–701.
- SAVIN, N. E. (1984): *Multiple Hypothesis Testing* Amsterdam: Elsevier Science Publishers BV, chap. Handbook of Econometrics, Volume II, pp. 827–879.
- SCHADY, N. R., AND M. C. ARAUJO (2006): “Cash, Conditions, School Enrollment, and Child Work: Evidence from a Randomized Experiment in Ecuador,” Unpublished manuscript.
- SCHULTZ, T. P. (2004): “School subsidies for the poor: Evaluating the Mexican PROGRESA poverty program,” *Journal of Development Economics*, 74(1), 199–250.
- SCHWARTZ, B., AND I. BHUSHAN (2004): “Reducing Inequity In The Provision Of Primary Health Care Services: Contracting In Cambodia,” mimeo.
- SMITH, J., AND P. TODD (2005): “Does Matching Overcome Lalondes Critique of Nonexperimental Estimators?,” *Journal of Econometrics*, 125(1-2), 305–353.
- THOMAS, D., E. FRANKENBERG, J. FRIEDMAN, J.-P. HABICHT, AND E. AL (2003): “Iron Deficiency and the Well Being of Older Adults: Early results from a randomized nutrition intervention,” Mimeo, UCLA.
- TODD, P. (2006): “Evaluating Social Programs with Endogenous Program Placement and Self Selection of the Treated,” forthcoming in the Handbook of Development Economics, Volume 4.
- TODD, P. E., AND K. I. WOLPIN (2006): “Ex Ante Evaluation of Social Programs,” Mimeo, University of Pennsylvania.
- VERMEERSCH, AND KREMER (2004): “School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation,” World Bank Policy Research Working Paper No. 3523.

WESTFALL, P., AND S. YOUNG (1993): *Resampling-based multiple testing: Examples and methods for p-value adjustment*. New York: Wiley.

WOOLDRIDGE (2002): “Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification,” *Portuguese Economic Journal*, 1, 117–139.

WOOLDRIDGE, J. M. (2004): “Cluster-Sample Methods in Applied Econometrics,” *American Economic Review*, 93(2), 133–138.

Table 1: Intra-class correlation, primary schools

Location	Subject	Estimate	Reference
Madagascar	Math+language	0.5	AGEPA data base
Busia, Kenya	Math+language	0.22	Miguel and Kremer (2004)
Udaipur, India	Math+language	0.23	Duflo and Hanna (2005)
Mumbai, India	Math+language	0.29	Banerjee et al. (2007)
Vadodara, India	Math+language	0.28	Banerjee et al. (2007)
Busia, Kenya	Math	0.62	Glewwe et al (2004)
Busia, Kenya	Language	0.43	Glewwe et al (2004)
Busia, Kenya	Science	0.35	Glewwe et al (2004)